

Channel Simulation and Coded Source Compression

Min-Hsiu Hsieh and Shun Watanabe

Abstract—This work establishes connection between channel simulation and coded source compression. First, we consider classical source coding with quantum side-information where the quantum side-information is observed by a helper and sent to the decoder via a classical channel. We derive a single-letter characterization of the achievable rate region for this problem. The direct part of our result is proved via the measurement compression theory by Winter, a quantum to classical channel simulation. Our result reveals that a helper's scheme that separately conducts a measurement and a compression is suboptimal, and the measurement compression is fundamentally needed to achieve the optimal rate region. We then study coded source compression in the fully quantum regime. We characterise the quantum resources involved in this problem, and derive a single-letter expression of the achievable rate region when entanglement assistance is available. The direct coding proof is based on a combination of two fundamental protocols, namely the quantum state merging protocol and the quantum reverse Shannon theorem. Our work hence resolves coded source compression in the quantum regime.

I. INTRODUCTION

Source coding normally refers to the information processing task that aims to reduce the redundancy exhibited when multiple copies of the same source are used. In establishing information theory, Shannon demonstrated a fundamental result that source coding can be done in a *lossless* fashion; namely, the recovered source will be an exact replica of the original one when the number of copies of the source goes to infinity [1]. If representing the source by a random variable X with output space \mathcal{X} and distribution p_X , lossless source coding is possible if and only if the compression rate R is above its Shannon entropy:

$$R \geq H(X), \quad (1)$$

where $H(X) := \sum_{x \in \mathcal{X}} -p_X(x) \log p_X(x)$.

Redundancy can also exist in the scenario in which multiple copies of the source are shared by two or more parties that are far apart. Compression in this particular setting is called *distributed* source coding, which has been proven to be extremely important in the internet era. The goal is to minimise the information sent by each party so that the decoder can still recover the source faithfully. Shannon's lossless source coding theorem can still be applied individually to each party.

This work was partially in Proceedings of IEEE Intl. Symp. Inf. Theory 2015, and in Proceedings of IEEE Information Theory Workshop 2015.

Min-Hsiu Hsieh is with the Centre for Quantum Computation & Intelligent Systems, Faculty of Engineering and Information Technology, University of Technology Sydney, Australia. Min-Hsiu Hsieh acknowledges the hospitality of the UTS-AMSS Joint Research Laboratory for Quantum Computation and Quantum Information Processing, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, China. (email: Min-Hsiu.Hsieh@uts.edu.au)

Shun Watanabe is with the Department of Computer and Information Sciences, Tokyo University of Agriculture and Technology, Japan. (email: shunwata@cc.tuat.ac.jp).

However, it is discovered that a better source coding strategy exists if the sources between different parties are correlated. Denote X and Y the sources held by the two distant parties, where the joint distribution is P_{XY} and the output spaces are \mathcal{X} and \mathcal{Y} , respectively. Slepian and Wolf showed that lossless distributed source coding is possible when the compression rates R_1 and R_2 for the two parties satisfy [2]:

$$R_1 \geq H(X|Y), \quad (2)$$

$$R_2 \geq H(Y|X), \quad (3)$$

$$R_1 + R_2 \geq H(XY), \quad (4)$$

where $H(X|Y)$ is the conditional Shannon entropy. This theorem is now called the classical Slepian-Wolf theorem [2]. In particular, when source Y is directly observed at the decoder, the problem is sometimes called source coding with (full) side-information.

Another commonly encountered scenario in a communication network is that a centralised server exists and its role is to coordinate all the information processing tasks, including the task of source coding, between the nodes in this network. Obviously, the role of the server is simply as a helper and it is not critical to reproduce the exact information communicated by the server. This slightly different scenario results in a completely different characterisation of the rate region, as observed by Wyner [3] and Ahlswede-Körner [4]. Consider that the receiver wants to recover the source X with the assistance of the server (that we will call a helper from now on) holding Y , where the distribution is P_{XY} . Wyner and Ahlswede-Körner showed that the optimal rate region for lossless source coding of X with a classical helper Y is the set of rate pairs (R_1, R_2) such that

$$R_1 \geq H(X|U), \quad (5)$$

$$R_2 \geq I(U; Y), \quad (6)$$

for some conditional distribution $p_{U|Y}(u|y)$, and $I(U; Y)$ is the classical mutual information between random variables U and Y . When there is no constraint on R_2 (i.e. R_2 can be as large as it can be), this problem reduces to source coding with (full) side-information.

The problem of source coding, when replacing classical sources with quantum sources, appears to be highly nontrivial in the first place¹. The first quantum source coding theorem was established by Schumacher [5], [6]. A quantum source ρ_A can be losslessly compressed and decompressed if and only if the rate R is above its von Neumann entropy²:

$$R \geq H(A)_\rho, \quad (7)$$

¹The quantum source coding result takes a much longer time to develop if one considers that quantum theory began to evolve in the mid-1920s.

²The subscript A is a label to which the quantum system ρ_A belongs.

where $H(A)_\rho := -\text{Tr } \rho_A \log \rho_A$.

Schumacher's quantum source coding theorem bears a close resemblance to its classical counterpart. One will naturally expect that the same will hold true for the distributed source coding problem in the quantum regime. Consider that Alice, who has the quantum system A of an entangled source ρ_{AB} , would like to merge her state to the distant party Bob. The quantum distributed source coding theorem (also known as quantum state merging) aims to answer the optimal rate R at which quantum states with density matrix ρ_A can be communicated to a party with quantum side information ρ_B faithfully. As it turns out, the optimal rate is given by the conditional von Neumann entropy $H(A|B)_\rho$, a quantum generalization of classical conditional Shannon entropy. While the quantum formula to the distributed source coding problem is also of the form of conditional entropy, this result has a much deeper and profound impact in the theory of quantum information as it marks a clear departure between classical and quantum information theory. It is rather perplexing that the rate R is quantified by the conditional entropy $H(A|B)_\rho$, which can be negative. This major piece of the puzzle was resolved with the interpretation that if the rate is negative, the state can be merged, and in addition, the two parties will gain $|H(A|B)_\rho|$ amount of entanglement for later quantum communication [8], [9], [10]. The distributed quantum source coding problem was later fully solved [11], [13] where the trade-off rate region between the quantum communication and the entanglement resource is derived. The result is now called the fully quantum Slepian-Wolf theorem (FQSW).

Source coding with hybrid classical-quantum systems ρ_{XB} with X representing a classical system and B a quantum state is also considered in quantum information theory, and one of our main results falls into this category. In [14], Devetak and Winter considered classical source coding with quantum side information at the decoder, and showed that the optimal rate R_1 is given by $H(X|B)_\rho$. This result can be regarded as a classical-quantum version of the source coding with (full) side-information.

In this work, we first revisit the classical coded source compression [3], [4]. In particular, we provide a simpler achievability proof based on the idea of channel simulation (Theorem 2). The proof indicates that channel simulation is a general subroutine employed between the helper and the decoder in the task of coded source compression.

Next, we consider classical source coding with a quantum helper. In our setup, the quantum side-information is observed by the helper, and the decoder will only have a classical description from the quantum helper. Although our problem can be regarded as a classical-quantum version of the classical helper problem studied in [3], [4], in contrast to its classical counterpart, our problem does not reduce to source coding with quantum side-information studied in [14] even if there is no constraint on rate R_2 . However, when the ensemble that constitutes the quantum side-information is commutative, our problem reduces to the classical helper problem [3], [4]. We completely characterize the rate region of the classical source coding with the quantum helper problem. In fact, the formulae describing the rate region (cf. Theorem 3) resembles

its classical counterpart (cf. (9) and (10)). However, the proof technique is very different due to the quantum nature of the helper. In particular, we use the measurement compression theory by Winter [23] in the direct coding theorem. One of interesting consequences of our result is that a helper's scheme that separately conducts a measurement and a compression is suboptimal; measurement compression is fundamentally needed to achieve the optimal rate region.

Next, we extend the classical distributed source coding problem [3], [4] and its classical-quantum generalisation to the fully quantum version; namely compression of a quantum source with the help of a quantum server. Moreover, we consider a general setting where entanglement assistance between sender-decoder and helper-decoder is available. Our direct coding proof combines two fundamental quantum protocols; the state merging protocol [8], [9] and the quantum reverse Shannon theorem [28]. The current progress of coded source compression is summarized in Table I. We remark that the quantum source compression with a classical helper is a very subtle task, which is left open even when the decoder has full classical-side information [14] (instead of partial side-information from the classical helper).

	Classical Source	Quantum Source
Classical Helper	Refs. [3], [4]	×
Quantum Helper	Theorem 3	Theorem 5

TABLE I
CODED SOURCE CODING IN CLASSICAL AND QUANTUM REGIMES

The quantum source compression with a quantum helper is treated in a different scenario in Ref. [9]. Overall there, classical communication is allowed from the helper to the receiver, and limited entanglement resource is considered. As a result, their formula requires regularization. In contrast, our result resorts to quantum reverse Shannon theorem, and has the appealing single-letter expression.

There are a huge amount of work devoted to both classical and quantum lossy source coding [15], [16], [17], [18], [19], [20]. We will restrict ourselves to only noiseless source coding in this work. However, as it turns out, channel simulation simplifies both rate distortion theory and coded source compression.

Notations. In this paper, we will use capital letters X, Y, Z, U etc. to denote classical random variables, and lower cases x, y, z, u to denote their realisations. We use $\mathcal{X}, \mathcal{Y}, \mathcal{Z}, \mathcal{U}$ to denote the sample spaces. We denote $x^n = x_1 x_2 \cdots x_n$.

A quantum state is a positive semi-definite matrix with trace equal to one. We will use ρ or σ to denote a quantum states in this paper. In case we need to specify which party the quantum state belongs to, we will use a subscript description ρ_A , meaning that the quantum system is held by Alice. Letting $\{|x\rangle\}_{x \in \mathcal{X}}$ be a set of orthonormal basis vectors, a classical-quantum state ρ_{XB} is written as

$$\rho_{XB} = \sum_x p_X(x) |x\rangle\langle x| \otimes \rho_x,$$

so that n copies of it is

$$\rho_{XB}^{\otimes n} = \sum_{x^n} p_X^{(n)}(x^n) |x^n\rangle\langle x^n| \otimes \rho_{x^n},$$

where we denote $\rho_{x^n} := \rho_{x_1} \otimes \cdots \otimes \rho_{x_n}$ for the sequence x^n . A positive-operator valued measure (POVM), $\Lambda = \{\Lambda_y\}_{y \in \mathcal{Y}}$, is a quantum measurement whose elements are non-negative self-adjoint operators on a Hilbert space so that $\sum_{y \in \mathcal{Y}} \Lambda_y = I$.

Various entropic quantities will be used in the paper. The von Neumann entropy of a quantum state ρ_A , where the subscript A represents the quantum state is held by Alice, is $H(A)_\rho = -\text{Tr}(\rho_A \log \rho_A)$. The conditional von Neumann entropy of system A conditioned on B of a bipartite state ρ_{AB} is $H(A|B)_\rho = H(AB)_\rho - H(B)_\rho$. The quantum mutual information between two systems A and B of ρ_{AB} is $I(A; B)_\rho = H(A)_\rho + H(B)_\rho - H(AB)_\rho$. The conditional quantum mutual information $I(A; B|C)_\rho = I(A; BC)_\rho - I(A; C)_\rho$.

We will also employ the framework of *Resource Inequality* (RI) [12], [13]. The RIs are a concise way of describing interconversion of resources in an information-processing task. Denote by $[qq]$ and $[q \rightarrow q]$ an ebit (maximally entangled pairs of qubits) and a noiseless qubit channel, respectively. Then a quantum channel \mathcal{N} that can faithfully transmit $Q(\mathcal{N})$ qubits per channel use with an unlimited amount of entanglement assistance can be symbolically represented as

$$\langle \mathcal{N} \rangle + \infty[qq] \geq Q(\mathcal{N})[q \rightarrow q],$$

where $\langle \mathcal{N} \rangle$ is an asymptotic noisy resource that corresponds to many independent uses, i.e. $\mathcal{N}^{\otimes n}$. Schumacher's noiseless source compression [5] can be similarly expressed

$$H(B)_\rho [q \rightarrow q] \geq \langle \rho_B \rangle,$$

which means that a rate of $H(B)_\rho$ noiseless qubits *asymptotically* is sufficient to represent the noisy quantum source ρ_B .

Sometimes, the RI only applies to the relative resource, $\langle \mathcal{N} : \rho \rangle$, which means that the asymptotic accuracy is achieved only when n uses of \mathcal{N} are fed an input of the form $\rho^{\otimes n}$. For detailed treatment of combining two RIs and cancellations of quantum resources, see Ref. [12].

This paper is organised as follows. In Sec II, we revisit classical source compression with a helper using the channel simulation idea. In Sec III, we formally define the problem of source coding with a quantum helper, and present the main result as well as its proof. In Sec IV, we treat the source coding with a helper in the fully quantum regime. We conclude in Sec V with open questions.

II. CLASSICAL CODED SOURCE COMPRESSION

Consider two classical channels $W_1 : \mathcal{X}_1 \rightarrow \mathcal{Y}_1$ and $W_2 : \mathcal{X}_2 \rightarrow \mathcal{Y}_2$. We define an (n, m, ϵ) channel simulation code consisting of the encoding and decoding pair $E_n : \mathcal{X}_2^m \rightarrow \mathcal{X}_1^n$, $D_n : \mathcal{Y}_1^n \rightarrow \mathcal{Y}_2^m$ that takes n uses of the first channel W_1 to simulate m instances of the channel output W_2 so that, $\forall x^m \in \mathcal{X}_2^m$

$$\|W_2^{\otimes m}(x^m) - D_n \circ W_1^{\otimes n} \circ E_n(x^m)\|_1 \leq \epsilon.$$

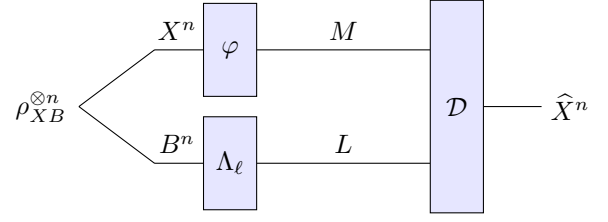


Fig. 1. Source Compression with a Quantum Helper.

The channel simulation capacity of the first channel W_1 needed to simulate the second channel W_2 is defined as:

$$C(W_1 \rightarrow W_2) = \lim_{\epsilon \rightarrow 0} \limsup_{n \rightarrow \infty} \left\{ \frac{m}{n} : \text{an } (n, m, \epsilon) \text{ code exists.} \right\} \quad (8)$$

The Shannon's noiseless channel coding theorem can be seen as a special case of $C(W_1 \rightarrow W_2)$ that W_2 is the identity channel id.

Another special case is the following classical reverse Shannon Theorem where W_1 is the identity channel id.

Theorem 1 (Classical Reverse Shannon Theorem [27]):

$$C(\text{id} \rightarrow N) = C(N).$$

where $C(N)$ is the Shannon capacity of the channel N .

Using the noiseless resource to simulate a noisy one seems a useless task to explore at first glance. However, it turns out that such a task can be used to simplify coded source compression conceptually (among others).

Theorem 2 (Classical source compression with a classical helper [3],

The optimal rate region for lossless source coding of X with a classical helper Y is the set of rate pairs (R_1, R_2) such that

$$R_1 \geq H(X|U), \quad (9)$$

$$R_2 \geq I(U; Y), \quad (10)$$

for some conditional distribution $p_{U|Y}(u|y)$, and $I(U; Y)$ is the classical mutual information between random variables U and Y .

With the help of Theorem 1, we can provide a simpler direct coding theorem for Theorem 2.

Proof: In the proof, the strategy that the classical helper does can be conceptually viewed as assisting the decoder to simulate the local channel $p_{U|Y}$. For n sufficiently large, the classical communication rate that the helper needs to send is $I(U; Y)$ and

$$\|p_{UX}^n - q_{UX}^n\|_1 \leq \epsilon,$$

where q_{UX}^n is the joint distribution induced by the simulation of $p_{U|Y}$. Thus, the full side information about U^n is possessed by the decoder, and the source compression with full side information can be carried on. Since the helper's local channel can be simulated at the decoder whose inaccuracy is at most ϵ , the overall error for classical source compression with a classical helper can be achieved with this additional ϵ error. ■

III. CLASSICAL SOURCE COMPRESSION WITH A QUANTUM HELPER

As shown in Figure 1, the protocol for classical source coding with a quantum helper involves two senders, Alice and Bob, and one receiver, Charlie. Initially Alice and Bob hold n copies of a classical-quantum state ρ_{XB} . In this case, Alice holds classical random variables X^n while Bob (being a helper) holds a quantum state ρ_{X^n} that is correlated with Alice's message. The goal is for the decoder Charlie to faithfully recover Alice's message when assisted by the quantum helper Bob.

We now proceed to formally define the coding procedure. We define an (n, ϵ) code for classical source compression with a quantum helper to consist of the following:

- Alice's encoding operation $\varphi : \mathcal{X}^n \rightarrow \mathcal{M}$, where $\mathcal{M} := \{1, 2, \dots, |\mathcal{M}|\}$ and $|\mathcal{M}| = 2^{nR_1}$;
- Bob's POVM $\Lambda = \{\Lambda_\ell\} : B^n \rightarrow \mathcal{L}$, where $\mathcal{L} := \{1, 2, \dots, |\mathcal{L}|\}$ and $|\mathcal{L}| = 2^{nR_2}$;
- Charlie's decoding operation $\mathcal{D} : \mathcal{M} \times \mathcal{L} \rightarrow \hat{\mathcal{X}}^n$

so that the error probability satisfies

$$\Pr\{X^n \neq \hat{X}^n\} \leq \epsilon. \quad (11)$$

A rate pair (R_1, R_2) is said to be *achievable* if for any $\epsilon, \delta > 0$ and all sufficiently large n , there exists an (n, ϵ) code with rates $R_1 + \delta$ and $R_2 + \delta$. The rate region is then defined as the collection of all achievable rate pairs. Our main result is the following theorem.

Theorem 3: Given is a classical-quantum source ρ_{XB} . The optimal rate region for lossless source coding of X with a quantum helper B is the set of rate pairs (R_1, R_2) such that

$$R_1 \geq H(X|U) \quad (12)$$

$$R_2 \geq I(U; B)_\sigma. \quad (13)$$

The state $\sigma_{UB}(\Lambda)$ resulting from Bob's application of the POVM $\Lambda = \{\Lambda_u\}_{u \in \mathcal{U}}$ is

$$\sigma_{UB}(\Lambda) = \sum_{u \in \mathcal{U}} p_U(u) |u\rangle\langle u| \otimes \rho_u \quad (14)$$

where

$$p_U(u) = \text{Tr}(\rho_B \Lambda_u) \quad (15)$$

$$\rho_u = \frac{1}{p_U(u)} [\sqrt{\rho_B} \Lambda_u \sqrt{\rho_B}]^* \quad (16)$$

$$\rho_B = \sum_x p_X(x) \rho_x. \quad (17)$$

where $*$ denotes complex conjugation in the standard basis. Furthermore, we can restrict the size of POVM as $|\mathcal{U}| \leq d_B^2$, where d_B is the dimension of Bob's system.

A typical shape of the rate region in Theorem 3 is described in Fig. 2. When there is no constraint on R_2 , rate R_1 can be decreased as small as

$$H(X|U^*) := \min_{\Lambda} H(X|U) \quad (18)$$

$$= H(X) - \max_{\Lambda} I(X; U) \quad (19)$$

$$= H(X) - I_{\text{acc}}, \quad (20)$$

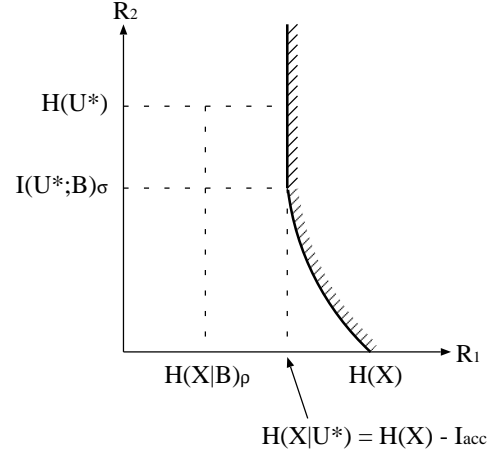


Fig. 2. A typical shape of the rate region in Theorem 3.

where I_{acc} is the accessible information for the ensemble $\{(p_X(x), \rho_x)\}_{x \in \mathcal{X}}$. Unless the ensemble commutes [7], the minimum rate $H(X|U^*)$ is larger than the rate $H(X|B)_\rho$, which is the optimal rate in the source coding with quantum side-information [14]. To achieve $R_1 = H(X|U^*)$, it suffices to have $R_2 \geq I(U^*; B)_\sigma$, which is smaller than $H(U^*)$ in general. This means that the following separation scheme is suboptimal: first conduct a measurement to get U^* and then compress U^* . For more detail, see the direct coding proof.

Converse:

Let $\varphi : \mathcal{X}^n \rightarrow \mathcal{M}$ be Alice's encoder, and let $\{\Lambda_\ell\}_{\ell \in \mathcal{L}}$ be Bob's measurement. Alice sends $M = \varphi(X^n)$ to the decoder, and Bob sends the measurement outcome L to the decoder. The Fano's inequality states that $H(X^n | M, L) \leq n\epsilon_n$ for some $\epsilon_n \rightarrow 0$ as $n \rightarrow \infty$.

First, we have the following bound:

$$\log |\mathcal{M}| \geq H(M) \quad (21)$$

$$\geq H(M|L) \quad (22)$$

$$\geq H(X^n|L) - H(X^n|M, L) \quad (23)$$

$$\stackrel{(a)}{\geq} H(X^n|L) - n\epsilon_n \quad (24)$$

$$\stackrel{(b)}{\geq} \sum_{t=1}^n H(X_t | X_{<t}, L) - n\epsilon_n \quad (25)$$

$$\stackrel{(c)}{=} \sum_{t=1}^n H(X_t | U_t) - n\epsilon_n \quad (26)$$

$$\stackrel{(d)}{=} nH(X_J | U_J, J) - n\epsilon_n, \quad (27)$$

where (a) follows from Fano's inequality: $H(X^n | M, L) \leq n\epsilon_n$ for some $\epsilon_n \rightarrow 0$ as $n \rightarrow \infty$; in (b), we use chain rule and denote $X_{<t} := (X_1, \dots, X_{t-1})$; in (c), we denote $U_t := (X_{<t}, L)$; in (d), we introduce a time-sharing random variable J that is uniformly distributed in the set $\{1, 2, \dots, n\}$.

Next, we have

$$\log |\mathcal{L}| \geq H(L) \quad (28)$$

$$\geq I(L; B^n) \quad (29)$$

$$= \sum_{t=1}^n I(L; B_t | B_{<t}) \quad (30)$$

$$= \sum_{t=1}^n I(L, B_{<t}; B_t) \quad (31)$$

$$\stackrel{(a)}{=} \sum_{t=1}^n I(L, B_{<t}, X_{<t}; B_t) \quad (32)$$

$$\geq \sum_{t=1}^n I(L, X_{<t}; B_t) \quad (33)$$

$$= \sum_{t=1}^n I(U_t; B_t) \quad (34)$$

where (a) follows from

$$I(X_{<t}; B_t | L, B_{<t}) \leq I(X_{<t}; B_t, B_{>t} | L, B_{<t}) \quad (35)$$

$$= H(X_{<t} | L, B_{<t}) - H(X_{<t} | L, B^n) \quad (36)$$

$$\leq H(X_{<t} | B_{<t}) - H(X_{<t} | L, B^n) \quad (37)$$

$$= H(X_{<t} | B_{<t}) - H(X_{<t} | B^n) \quad (38)$$

$$= H(X_{<t} | B_{<t}) - H(X_{<t} | B_{<t}) \quad (39)$$

$$= 0. \quad (40)$$

Following from Eq. (34), we can again introduce a time-sharing random variable J that is uniformly distributed in the set $\{1, 2, \dots, n\}$,

$$\sum_{t=1}^n I(U_t; B_t) = n \sum_{t=1}^n I(U_t; B_t | J = t) \quad (41)$$

$$= n I(U_J; B_J | J) \quad (42)$$

$$= n I(U_J; B_J) \quad (43)$$

where the last equality follows because $I(J; B_J) = 0$. To get single-letter formula, define $X = X_J$, $B = B_J$, and $U = (U_J, J)$ and let $n \rightarrow \infty$:

$$R_1 = \frac{1}{n} \log |\mathcal{M}| \geq H(X|U) \quad (44)$$

$$R_2 = \frac{1}{n} \log |\mathcal{L}| \geq I(U; B). \quad (45)$$

Here, we note that the distribution of $U_t = (L, X_{<t})$ can be written as

$$p_{X_{<t}L}(x_{<t}, \ell) = \left(\prod_{i<t} p_X(x_i) \right) \times \text{Tr} \left[\left\{ \left(\bigotimes_{i<t} \rho_{x_i} \right) \otimes \rho_{B_t} \otimes \left(\bigotimes_{i>t} \rho_{B_i} \right) \right\} \Lambda_\ell \right]. \quad (46)$$

Thus, we can get U_t as a measurement outcome of B_t by first generating $X_{<t}$, then by appending $\bigotimes_{i<t} \rho_{x_i}$ and $\bigotimes_{i>t} \rho_{B_i}$ to ancillae systems, and finally by conducting the measurement $\{\Lambda_\ell\}_{\ell \in \mathcal{L}}$.

Finally, the bound on $|\mathcal{U}|$ can be proved via Carathéodory's theorem (cf. [26, Appendix C]). ■

Direct Coding Theorem: Fix a POVM measurement $\Lambda = \{\Lambda_u\}_{u \in \mathcal{U}}$. It induces a conditional probability $p_{U|X}(u|x) = \text{Tr}[\Lambda_u \rho_x]$, and joint probability distribution

$$P_{XU}(x, u) = p_X(x) p_{U|X}(u|x). \quad (47)$$

The crucial observation is the application of Winter's measurement compression theory [23].

Theorem 4 (Measurement compression theorem [23], [24]): Let ρ_A be a source state and Λ a POVM to simulate on this state. A protocol for a faithful simulation of the POVM is achievable with classical communication rate R and common randomness rate S if and only if the following set of inequalities hold

$$R \geq I(X; R), \quad R + S \geq H(X), \quad (48)$$

where the entropies are with respect to a state of the following form:

$$\sum_x |x\rangle \langle x|^X \otimes \text{Tr}_A \{ (I^R \otimes \Lambda_x^A) \phi^{RA} \}, \quad (49)$$

and ϕ^{RA} is some purification of the state ρ_A .

Let K be a random variable on \mathcal{K} , which describes the common randomness shared between Alice and Bob. Let $\{\tilde{\Lambda}_{u^n}^{(k)}\}_{u^n \in \mathcal{U}^n}$ be collection of POVMs. Let

$$Q_{X\tilde{U}}^n(x^n, u^n) := P_X^{(n)}(x^n) \sum_{k \in \mathcal{K}} \frac{1}{|\mathcal{K}|} \text{Tr}[\rho_{x^n} \tilde{\Lambda}_{u^n}^{(k)}], \quad (50)$$

where $P_X^{(n)}(x^n) := P_X(x_1) \times \dots \times P_X(x_n)$. The faithful simulation of n copies of POVM $\Lambda := \{\Lambda_u\}_{u \in \mathcal{U}}$, i.e. $\Lambda^{\otimes n}$, implies that for any $\epsilon > 0$, there exists n sufficiently large, such that there exist POVMs $\{\tilde{\Lambda}^{(k)}\}$, where $\tilde{\Lambda}^{(k)} := \{\tilde{\Lambda}_{u^n}^{(k)}\}_{u^n \in \mathcal{U}^n}$, with

$$\frac{1}{2} \|P_{XU}^{(n)} - Q_{X\tilde{U}}^n\|_1 \leq \epsilon. \quad (51)$$

Coding Strategy:

Alice and Bob shared n copies of the state ρ_{XB} , and assume that Bob performs measurement $\Lambda^{\otimes n} : B^{\otimes n} \rightarrow \mathcal{U}^n$ on his quantum system whose outcome is sent to the decoder to assist decoding Alice's message. Bob's measurement on each copy of ρ_{XB} will induce the probability distribution P_{XU} according to (47). Apparently, if Bob sends the full measurement outcomes to Charlie (say $nH(U)$ bits), then Charlie can successfully decode X^n simply from Slepian-Wolf Theorem. The next strategy is to make use of classical result since after Bob's measurement, Alice and Bob become fully classical with joint distribution P_{XU} . Therefore, the minimum rate for Bob is $I(V; U)$ (w.r.t. some conditional distribution $p_{V|U}$). However, there is a non-trivial quantum coding strategy. Detail follows.

Bob's coding. Instead of the measurement Λ performed on Bob's system ρ_B and coding w.r.t. the classical channel $p_{V|U}(v|u)$, the decoder Charlie can directly simulate the measurement outcome U using Winter's measurement compression theorem [23], [24]. Denote Bob's classical communication rate $R_2 = \frac{1}{n} \max_{k \in \mathcal{K}} |\tilde{\Lambda}^{(k)}|$. Then Theorem 4 promises

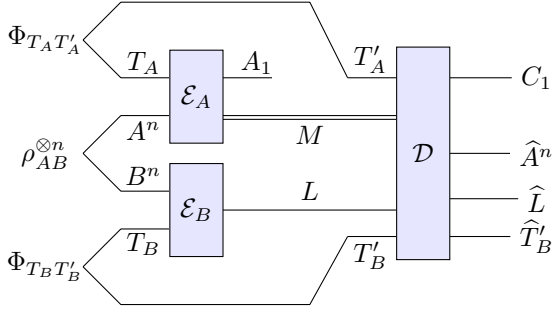


Fig. 3. Fully Quantum Source Compression with a Quantum Helper.

that by sending $R_2 \geq I(U; B)$ from Bob to the decoder Charlie, Charlie will have a local copy \tilde{U}^n and the distribution between Alice and Charlie $Q_{X\tilde{U}}^n$ will satisfy (51).

Alice's coding. Now Alice's strategy is very simple since Charlie has had \tilde{U}^n . She just uses the Slepian-Wolf coding strategy as if she starts with the distribution P_{XU} with Charlie. In fact, it is well known (cf. [21]) that there exists an encoder $\varphi: \mathcal{X}^n \rightarrow \mathcal{M}$ and a decoder $\mathcal{D}: \mathcal{M} \times \mathcal{U}^n \rightarrow \mathcal{X}^n$ such that $|\mathcal{M}| = 2^{n(H(X|U)+\delta)}$ and

$$P_{XU}^{(n)}(\mathcal{A}^c) \leq \epsilon \quad (52)$$

for sufficiently large n , where

$$\mathcal{A} := \{(x^n, u^n) \in \mathcal{X}^n \times \mathcal{U}^n : \mathcal{D}(\varphi(x^n), u^n) = x^n\} \quad (53)$$

is the set of correctly decodable pairs.

Now, suppose that Alice and Bob use the same code for the simulated distribution $Q_{X\tilde{U}}^n$. Then, by the definition of the variational distance and (51), we have

$$Q_{X\tilde{U}}^n(\mathcal{A}^c) \leq P_{XU}^{(n)}(\mathcal{A}^c) + \epsilon. \quad (54)$$

Thus, if we can find a good code for $P_{XU}^{(n)}$, we can also use that code for $Q_{X\tilde{U}}^n$ for sufficiently large n .

Derandomization. The standard derandomization technique works here. If the random coding strategy works fine on average, then there exists one realisation works fine too. Since the distribution $Q_{X\tilde{U}}^n = \frac{1}{|\mathcal{K}|} \sum_{k \in \mathcal{K}} Q_{X\tilde{U}|k}^n$ and

$$\sum_k \frac{1}{|\mathcal{K}|} Q_{X\tilde{U}|k}^n(\mathcal{A}^c) = Q_{X\tilde{U}}^n(\mathcal{A}^c) \leq P_{XU}^{(n)}(\mathcal{A}^c) + \epsilon. \quad (55)$$

Thus, there exists one $k \in \mathcal{K}$ so that $Q_{X\tilde{U}|k}^n(\mathcal{A}^c)$ is small. ■

IV. FULLY QUANTUM SOURCE COMPRESSION WITH A QUANTUM HELPER

As shown in Figure 3, the protocol for fully quantum source coding with a quantum helper involves two senders, Alice and Bob, and one receiver, Charlie. Initially Alice and Bob hold n copies of a bipartite quantum state ρ_{AB} , where Alice holds quantum systems $A^n := A_1 \cdots A_n$ while Bob (being a quantum helper) holds quantum systems $B^n = B_1 \cdots B_n$. Moreover, there are pre-shared entangled states $|\Phi_{T_A T'_A}\rangle$ between Alice and Charlie, and pre-shared entangled states $|\Phi_{T_B T'_B}\rangle$ between Bob and Charlie. The goal is for the

decoder Charlie to faithfully recover Alice's quantum state $\rho_{A^n} = \text{Tr} \rho_{AB}^{\otimes n}$ when assisted by the quantum helper Bob.

We now proceed to formally define the coding procedure. Let ψ_{ABR} be a purification of ρ_{AB} . We define an (n, ϵ) code for fully quantum source compression with a quantum helper to consist of the following:

- Alice's encoding operation $\mathcal{E}_A: T_A A^n \rightarrow A_1 M$, where A_1 is a quantum system and M is a classical system; Alice only sends M to Charlie;
- Bob's encoding operation $\mathcal{E}_B: T_B B^n \rightarrow L$, where L is a quantum system of dimension $|L| = 2^{nR_2}$; Bob sends the quantum system to Charlie;
- Charlie's decoding operation $\mathcal{D}: M L T'_A T'_B \rightarrow C_1 \hat{A}^n \hat{L} \hat{T}'_B$ that produces

$$\omega_{A_1 C_1 \hat{A}^n \hat{L} R^n \hat{T}'_B} = I_{A_1} \otimes \mathcal{D}(\sigma_{A_1 M L R^n T'_A T'_B})$$

where

$$\sigma_{A_1 M L R^n T'_A T'_B} = \mathcal{E}_A \otimes \mathcal{E}_B \otimes I_{R^n} (\psi_{ABR}^{\otimes n} \otimes \Phi_{T_A T'_A} \otimes \Phi_{T_B T'_B});$$

so that the final state satisfies

$$\|\omega_{A_1 C_1 \hat{A}^n \hat{L} R^n \hat{T}'_B} - \Phi_{A_1 C_1} \otimes \rho_{A^n L R^n T'_B}\|_1 \leq \epsilon, \quad (56)$$

where $|\Phi_{A_1 C_1}\rangle$ is a maximally entangled state.

Let $R_1 = \log |T_A| - \log |A_1|$. A rate pair (R_1, R_2) is said to be *achievable* if for any $\epsilon, \delta > 0$ and all sufficiently large n , there exists an (n, ϵ) code with rates $R_1 + \delta$ and $R_2 + \delta$. The rate region is then defined as the collection of all achievable rate pairs. Our main result is the following theorem.

Theorem 5: Given is a bipartite quantum source $\rho_{AB} = \text{Tr}_R \psi_{ABR}$. The optimal rate region for lossless source coding of A with a quantum helper B is the set of rate pairs (R_1, R_2) such that

$$R_1 \geq H(A|C)_\phi \quad (57)$$

$$R_2 \geq \frac{1}{2} I(RA; C)_\phi. \quad (58)$$

The state ϕ_{ACER} resulting from Bob's application of some CPTP map $\mathcal{E}_{B \rightarrow C}$ is

$$|\phi_{ACER}\rangle = I_{RA} \otimes U_{B \rightarrow CE}^\mathcal{E} |\psi_{ABR}^p\rangle. \quad (59)$$

A. Direct part

1) Relevant quantum protocols: Given a bipartite state ρ_{AB} whose purification is $|\psi_{ABR}\rangle$, the state merging protocol [8], [9], [10] is the information-processing task of distributing A -part of the system that originally belongs to Alice to the distant Bob without altering the joint state. Moreover, Alice and Bob have access to pre-shared entanglement and their goal is to minimise the number of EPR pairs consumed during the protocol. The state merging can be efficiently expressed as the following RI:

$$\langle \psi_{A|B|R} \rangle + I(A; R)_\psi [c \rightarrow c] + H(A|B)_\psi [qq] \geq \langle \psi_{|AB|R} \rangle \quad (60)$$

where the notation $\psi_{A|B|R}$ denotes the state is originally shared between three distant parties Alice, Bob, and Eve, while $\psi_{|AB|R}$ means that the system A is now together with system B . This protocol involves classical communication; however,

for the purpose of this paper, quantum resources are much more valuable and classical communication is considered to be free. As a result, the state merging protocol either consumes EPR pairs with rate $H(A|B)_\psi$ when this quantity is positive, or generates $H(A|B)_\psi$ rate of EPR pairs for later uses, if $H(A|B)_\psi$ is negative, after the transmission of the system A to B .

The state merging protocol gives the first operational interpretation to the conditional von Neumann entropy. Most importantly, it provides an answer to the long-standing puzzle—the conditional von Neumann entropy could be negative, a situation that has no classical correspondence.

The fully quantum Slepian-Wolf (FQSW) protocol [11], [13] can be considered as the coherent version of the state merging protocol. It can be described as

$$\langle \psi_{A|B|R} \rangle + \frac{1}{2}I(A; R)_\psi[q \rightarrow q] \geq \frac{1}{2}I(A; B)_\psi[qq] + \langle \psi_{|AB|R} \rangle. \quad (61)$$

It is a simple exercise to show, via the resource inequalities, that the state merging protocol can be obtained by combining teleportation with the FQSW protocol [11], [13]. Moreover, the FQSW protocol can be transformed into the a version of the quantum reverse Shannon theorems (QRST) that involves entanglement assistance [11].

The quantum reverse Shannon theorem (QRST) addresses a fundamental task that asks, given a quantum channel \mathcal{N} , how much quantum communication is required from Alice to Bob so that the channel \mathcal{N} can be simulated. There are variants of the QRSTs depending on whether entanglement or feedback is allowed in the simulation (see [28, Theorem 3]). The QRST protocol has become a powerful tool in quantum information theory. It can be used to establish a strong converse to the entanglement-assisted capacity theorem. Moreover, it can also be used to establish quantum rate distortion theorems [18], [19], [20].

In this paper, we will use the QRST with entanglement assistance.

Theorem 6 (Quantum Reverse Shannon Theorem): Let \mathcal{N} be a quantum channel from A to B so that its isometry $U_{A \rightarrow BE}^\mathcal{N}$ results in the following tripartite state when inputting ρ_A :

$$|\psi_{RBE}\rangle = U_{A \rightarrow BE}^\mathcal{N} |\psi_{RA}^\rho\rangle,$$

where $\text{Tr}_R |\psi_{RA}^\rho\rangle\langle\psi_{RA}^\rho| = \rho_A$. Then with sufficient amount of pre-shared entanglement, the channel \mathcal{N} with input ρ_A can be simulated with quantum communication rate $\frac{1}{2}I(R; B)_\psi$:

$$\frac{1}{2}I(R; B)_\psi[q \rightarrow q] + \frac{1}{2}I(E; B)_\psi[qq] \geq \langle \mathcal{N} : \rho_A \rangle. \quad (62)$$

Proof:

We use the channel simulation method. For any local channel $\mathcal{E}_{B \rightarrow C}$ performed by the quantum helper B on his half of bipartite state ρ_{AB} , it can be simulated by the decoder using the quantum reverse Shannon theorem (QRST) (Theorem 6):

$$\frac{1}{2}I(RA; C)_\phi[q \rightarrow q] + \frac{1}{2}I(E; C)_\phi[qq] \geq \langle \mathcal{E} : \rho_B \rangle, \quad (63)$$

where

$$|\phi_{ACER}\rangle = I_R \otimes U_{B \rightarrow CE}^\mathcal{E} |\psi_{ABR}^\rho\rangle.$$

In other words, by using the pre-shared entanglement between the helper and the decoder with rate $\frac{1}{2}I(E; C)_\phi$ and sending quantum message from the helper to the decoder with rate $\frac{1}{2}I(RA; C)_\phi$, the decoder can simulate the quantum state $\mathcal{E}(\rho_B)$ locally with error goes to zero in the asymptotic sense.

Alice's coding: Once the decoder has the system C , then Alice and the decoder start the state merging protocol, using the pre-shared entanglement with rate $H(A|C)_\phi$. ■

B. Converse part

Here, we refer to Figure 3 for corresponding labels used in the converse proof. To bound $R_1 = \log |T_A| - \log |A_1|$, we follow steps in the converse proof of the state merging protocol [9] and have

$$R_1 \gtrsim H(A^n | LT'_B) \quad (64)$$

$$= \sum_{i=1}^n H(A_i | LT'_B A_{<i}) \quad (65)$$

$$= \sum_{i=1}^n H(A_i | U_i) \quad (66)$$

$$= nH(A_T | U_T T) \quad (67)$$

$$= nH(A|C), \quad (68)$$

where we set $U_i := (L, T'_B, A_{<i})$ and in the last equality, we relabel $A = A_T$ and $C := (U_T, T)$.

To bound the quantum communication rate $R_2 = \log |L|$, we follow steps in the converse proof of the entanglement-assisted quantum rate-distortion theorem (see equation (21) in [19]):

$$2R_2 \geq I(LT'_B; R^n A^n) \quad (69)$$

$$= \sum_{i=1}^n I(LT'_B; R_i A_i | R_{<i} A_{<i}) \quad (70)$$

$$= \sum_{i=1}^n [I(LT'_B R_{<i} A_{<i}; R_i A_i) - I(R_{<i} A_{<i}; R_i A_i)] \quad (71)$$

$$= \sum_{i=1}^n I(LT'_B R_{<i} A_{<i}; R_i A_i) \quad (72)$$

$$\geq \sum_{i=1}^n I(LT'_B A_{<i}; R_i A_i) \quad (73)$$

$$= \sum_{i=1}^n I(U_i; R_i A_i) \quad (74)$$

$$= nI(U_T; R_T A_T | T) \quad (75)$$

$$= nI(U_T T; R_T A_T) \quad (76)$$

$$= nI(C; RA). \quad (77)$$

Note that U_i can be generated from B_i via Bob's local CPTP. In fact, Bob can first append the maximally entangled states (T_B, T'_B) , systems $(A_{<i}, B_{<i})$, and $B_{>i}$. Then, he can perform \mathcal{E}_B , and get $U_i = (L, T'_B, A_{<i})$.

V. CONCLUSION AND DISCUSSION

We considered the problem of compression of a classical source with a quantum helper. We completely characterised its rate region and showed that the capacity formula does not require regularisation, which is not common in the quantum setting. While the expressions for the rate region are similar to the classical result in [3], [4], [22], it requires very different proof technique. To prove the achievability, we employed a powerful theorem, measurement compression theorem [23], that can decompose quantum measurement. A similar approach was recently applied to derive a non-asymptotic bound on the classical helper problem [25].

The rate region in our Theorem 5 bears a close resemblance to its classical counterpart. Our result also shows a helper's strategy of simply compressing the side information $H(C)_\phi$ and sending it to the decoder is sub-optimal with entanglement assistance. Recall the following identity:

$$H(C)_\phi = \frac{1}{2}I(C; E)_\phi + \frac{1}{2}I(C; RA)_\phi,$$

where the state $|\phi_{ACER}\rangle$ is given in (59). The QRST protocol allows us to cleverly divide the amount of quantum communication required for lossless transmission of system C to the decoder into pre-shared entanglement with rate $\frac{1}{2}I(C; E)_\phi$ and quantum communication with rate $\frac{1}{2}I(C; RA)_\phi$.

We will like to point out that the definition of the fully quantum source compression with a quantum helper requires to explicitly include additional quantum systems LT'_B (see Eq. (56)) for a technical purpose. The reason behind this is because when the quantum state merging is performed, the target systems to which the quantum state is merged needs to be specified. We believe that the inclusion of these additional systems in the definition is inevitable, and it signals a fundamental difference between the fully quantum source compression with a quantum helper and its classical counterpart.

Note that it is possible to replace the state merging protocol with the FQSW protocol, and derive an alternative theorem for quantum source compression with a quantum helper. It is also possible to consider the same problem without entanglement assistance between the helper and the decoder. These extensions will be treated in the future.

Finally, in the classical source coding with a helper problem, it is possible to bound the dimension for the helper's output system. However, such a result is not unknown to be possible in the quantum regime.

ACKNOWLEDGEMENTS

MH is supported by an ARC Future Fellowship under Grant FT140100574. SW was supported in part by JSPS Postdoctoral Fellowships for Research Abroad.

REFERENCES

- [1] C. E. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, vol. 27, pp. 379–423, 623–656, 1948.
- [2] D. Slepian and J. K. Wolf, "Noiseless coding of correlated information sources," *IEEE Trans. Inform. Theory*, vol. 19, no. 4, pp. 471–480, 1973.
- [3] A. D. Wyner, "On source coding with side information at the decoder," *IEEE Trans. Inform. Theory*, vol. 21, no. 3, pp. 294–300, 1975.
- [4] R. Ahlswede and J. Körner, "Source coding with side information and a converse for the degraded broadcast channel," *IEEE Trans. Inform. Theory*, vol. 21, no. 6, pp. 629–637, 1975.
- [5] B. Schumacher, "Quantum coding," *Phys. Rev. A*, vol. 51, no. 4, pp. 2738–2747, Apr. 1995.
- [6] R. Jozsa and B. Schumacher, "A New Proof of the Quantum Noiseless Coding Theorem," *J. of Modern Optics*, vol. 41, no. 12, pp. 2343–2349, Dec. 1994.
- [7] P. Hayden, R. Jozsa, D. Petz, and A. Winter, "Structure of States Which Satisfy Strong Subadditivity of Quantum Entropy with Equality," *Communications in Mathematical Physics*, vol. 246, no. 2, pp. 359–374, Feb. 2004.
- [8] M. Horodecki, J. Oppenheim, and A. Winter, "Partial quantum information," *Nature*, vol. 436, no. 7051, pp. 673–676, Aug. 2005.
- [9] M. Horodecki, J. Oppenheim, and A. Winter, "Quantum State Merging and Negative Information," *Communications in Mathematical Physics*, vol. 269, no. 1, pp. 107–136, Oct. 2006.
- [10] F. Dupuis, M. Berta, J. Wullschleger, and R. Renner, "One-Shot Decoupling," *Communications in Mathematical Physics*, vol. 328, no. 1, pp. 251–284, Mar. 2014.
- [11] A. Abeyesinghe, I. Devetak, P. M. Hayden, and A. Winter, "The mother of all protocols: restructuring quantum information's family tree," *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 465, no. 2108, pp. 2537–2563, Jun. 2009.
- [12] I. Devetak, A. W. Harrow, and A. Winter, "A Resource Framework for Quantum Shannon Theory," *IEEE Trans. Inform. Theory*, vol. 54, no. 10, pp. 4587–4618, Oct. 2008.
- [13] N. Datta and M.-H. Hsieh, "The apex of the family tree of protocols: optimal rates and resource inequalities," *New J. Phys.*, vol. 13, no. 9, p. 093042, 2011.
- [14] I. Devetak and A. Winter, "Classical data compression with quantum side information," *Phys. Rev. A*, vol. 68, no. 4, Oct. 2003.
- [15] C. E. Shannon, "Coding theorems for a discrete source with a fidelity criterion," *IRE Nat. Conv. Rec.*, vol. 4, pp. 142–163, 1959.
- [16] T. Berger, *Rate Distortion Theory: A Mathematical Basis for Data Compression*. Englewood Cliffs, NJ: Prentice Hall, 1971.
- [17] I. Devetak and T. Berger, "Quantum rate-distortion theory for memoryless sources," *IEEE Trans. Inform. Theory*, vol. 48, no. 6, pp. 1580–1589, 2002.
- [18] N. Datta, M.-H. Hsieh, and M. M. Wilde, "Quantum Rate Distortion, Reverse Shannon Theorems, and Source-Channel Separation," *IEEE Trans. Inform. Theory*, vol. 59, no. 1, pp. 615–629, 2013.
- [19] M. M. Wilde, N. Datta, M.-H. Hsieh, and A. Winter, "Quantum Rate-Distortion Coding With Auxiliary Resources," *IEEE Trans. Inform. Theory*, vol. 59, no. 10, pp. 6755–6773, 2013.
- [20] N. Datta, M.-H. Hsieh, M. M. Wilde, and A. Winter, "Quantum-to-classical rate distortion coding," *J. Math. Phys.*, vol. 54, no. 4, p. 042201, 2013.
- [21] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. Wiley, New York, 1991.
- [22] A. El Gamal and Y.-H. Kim, *Network information theory*. Cambridge University Press, 2011.
- [23] A. Winter, "Extrinsic" and "Intrinsic" Data in Quantum Measurements: Asymptotic Convex Decomposition of Positive Operator Valued Measures. *Communications in Mathematical Physics*, 244(1), 157–185, 2004.
- [24] M. M. Wilde, P. M. Hayden, F. Buscemi, and M.-H. Hsieh, "The information-theoretic costs of simulating quantum measurements," *Journal of Physics A: Mathematical and Theoretical*, vol. 45, no. 45, pp. 453001, Nov. 2012.
- [25] S. Watanabe, S. Kuzuoka, and V. Y. F. Tan, "Non-Asymptotic and Second-Order Achievability Bounds for Source Coding With Side-Information," in *Proc. 2013 IEEE International Symposium on Information Theory*, pp. 3055–3059.
- [26] I. Devetak and A. Winter, "Distillation of secret key and entanglement from quantum states," in *Proc. of The Royal Society A*, vol. 461, pp. 207–235, Jan. 2005.
- [27] Bennett, Charles, Peter W Shor, John A Smolin, and A V Thapliyal. "Entanglement-Assisted Capacity of a Quantum Channel and the Reverse Shannon Theorem." *IEEE Transactions on Information Theory* 48, no. 10 (October 2002): 2637–2655. doi:10.1109/TIT.2002.802612.
- [28] C. H. Bennett, I. Devetak, A. W. Harrow, P. W. Shor, and A. Winter. The Quantum Reverse Shannon Theorem and Resource Tradeoffs for Simulating Quantum Channels. *IEEE Trans. Inform. Theory*, vol. 60, no. 5, pp. 2926–2959, 2014.
- [29] M.-H. Hsieh and S. Watanabe. Source Compression with a Quantum Helper. arXiv:1501.04366, 2015 (accepted for publication in ISIT 2015).